Child Income as an Insurance Mechanism. Consequences for the Health-Education Relationship

Renaud Bourlès, Bruno Ventelou, Maame Esi Woode

Working Papers / Documents de travail

Child Income as an Insurance Mechanism

Consequences for the Health-Education Relationship

Renaud Bourlès
Bruno Ventelou
Maame Esi Woode

WP 2012 - Nr 05

Child Income as an Insurance Mechanism

Consequences for the Health-Education Relationship

Renaud Bourlès[a,b,*], Bruno Ventelou[a,c*]and Maame Esi Woode[a†]

[a] *Aix-Marseille School of Economics*

[b] *Ecole Centrale Marseille*

[c] *INSERM SE4S*

**E-mail addresses*: renaud.bourles@univmed.fr, bruno.ventelou@inserm.fr

†*Corresponding Author: GREQAM, Centre de Vieille Charité, 2, Rue de la Charité, 13236 Marseille Cedex 02, France. Tel: +33491114073. Cell: +33629601092. Fax: +33491900227. E-mail: mwoode@gmail.com*

Abstract

This paper analyzes the relationships between HIV/AIDS and education taking into account the appropriative nature of child income. We first build a simple theoretical model linking parental health risk, educational choice and appropriation of future children's income. We show that considering (remittances from) child's income as an insurance asset can reverse the usual negative relationship between disease prevalence and educational investment. This prediction is tested on data compiled from the Demographic and Health Survey (DHS) database for 17 Sub-Sahara African (SSA) countries between the years 2003 to 2010 for children aged between 6 and 22-years-old. To account for the hierarchical nature of the data we employ a multilevel analysis. We find that, in general, the impact of community HIV prevalence on school enrollment is insignificant. Once the data is split to account for differences in appropriation, the effect of community prevalence becomes positive and sometimes significant for highly appropriable groups (rural, girls) and remains either negative for the rest.

Keywords: *Health risk; Education; Insurance mechanism; Remittance.*

JEL Classification Numbers: I15, I25.

1 Introduction

Education has been proven to be key in developping countries, regarding both macroeconomic (growth) and microeconomic (consumption smoothing, risk management) issues. Within risk-sharing systems, education can act as a way to smooth consumption and to reduce the risk of wealth loss induced by health risk. It is now well documented that health status and in particular HIV status has an important (negative) impact on child education. This is particularly true in developing countries where this negative relationship can lead to macroeconomic issues, such as poverty trap. However, if we focus on the impact of health risk (which can be proxied by prevalence) rather than health status, the impact on education seems less obvious. We therefore propose here to enrich the analysis of the link between HIV and education by focusing – both theoretically and empirically – on health risk (rather than status) and on the insurance role of education.

This paper refers to three fields of economic literature: i) the relationship between epidemics and economic development at the micro level, i.e. how a health shock may jeopardize a household's future capacity to save, accumulate and generate future income (see Kawabata et al. 2002, Wagstaff and Doorslaer 2003, Xu et al. 2003 and Flores et al. 2008[1]; Fortson (2011) already demonstrates the impact of HIV/AIDS on children's education), ii) child labor and child income appropriation by parents (for example Basu and Pham 1998 and Schoonbroodt and Tertilt 2010), iii) educational choice as old-age insurance in the absence of social security, the most common situation in a lot of African countries (Ehrlich and Lui 1991 for example studied the effect of changing mortality in an overlapping-generations model in which children provide old-age support for their parents).

There are two ways of insuring for old age: educating one's children or having a lot of children. Investment in children, both quantitatively and qualitatively is an asset for parents, seen in developing countries as a substitute for social security (see for example Nugent 1985, Michel and Wigniole 2007 and Ehrlich and Lui 1991), especially in old age, which is often characterized by the absence of labor income and the presence of elevated health-risk. However, the fundamental nature of this insurance asset is that it is mobilized differently depending on the realization of a risk: actual

[1] The welfare effects of HIV can be approximated through the share of the household budget absorbed by HIV related medical expenditures, (see for example Wagstaff and Doorslaer 2003 and Xu et al. 2003). When health expenditures represent a large share of the healthcare the situation is defined as "catastrophic" as this implies either a substantial sacrifice of the household current consumption, possibly including basic needs, or the mobilization of other resources (such as savings, assets, credit and transfers from friends and family).

repayments are contingent on bad event. Starting from this point, we add to the idea of investment in children as an old-age insurance the notion of *changing appropriation*. This refers to the return on investment in a child's education being dependent on a variety of characteristics of the child and it's household. Moreover, the intensity with which parents invoke their "rights" on the income of their descendants varies depending on their health status.

This is not only inspired by the need to draw a parallel with the classic insurance framework, but can be connected to the notion of "child income appropriation" (see Schoonbroodt and Tertilt 2010): in a lot of developing countries, the property rights on labor-income are shared between the different members of a household and, at least partially, submitted to discretionary re-appropriation decisions coming from the household community. Bazen and Salmon (2010) and McIntyre et al. (2006) showed that "child-labor" and "premature removal from school" are connected with health shocks, suggesting that child income indeed plays the role of an insurance asset, with repayments contingent on a bad event. The recent paper of Maccini and Yang (2009) clearly follows the same direction, but links schooling decisions to other types of shock (rainfall shocks) and focuses on the asymmetric effect between girls and boys in the household (as in Duflo 2000). More generally, if we extend the analysis to fields outside schooling choice, numerous authors, as Fafchamps (1992), Harrower and Hoddinott (2005) and Park (2006), consider that shocks are insured through risk sharing networks (not only household) and that remittances of labor-income act as contingent repayments in case of negative shocks (see Gertler and Gruber 2002, Conroy et al. 2007). These intra-community transfers are more or less compulsory, due to the need of reciprocity in the co-insurance system of the network (Fafchamps 1992), or due to social punishment (Rapoport and Docquier 2006).

In the following, we first present a theoretical model to obtain testable predictions on the relationship between health shock (on adults) and child education considering child education as an insurance asset. We present the property of a "changing appropriation" in case of bad event. We reinterpret the impact of the health shock on educational choice of households in a totally new manner which allows to deal with an empirical "puzzle". Indeed, our empirical findings are not compatible with the existing view that a health shock necessarily has a negative effect on the education of children.

2 Theoretical Model

In this section we build a simple theoretical model to highlight the importance of child income appropriation for the relationship between education and disease prevalence. While the literature examines the quantitative choice (i.e. fertility) in respond to health shocks (see Becker and Barro 1988 for a survey), we focus on the qualitative decision on education enrollment, modeling here a two-person household. We consider the problem of a representative household composed of a head (the parent) and a child in a two-period model[2].

In the first period (activity period), the parent has a secure revenue, ω, and chooses the proportion of time they educate their child, t. When not enrolled in the educational system, that is in a proportion $(1-t)$ of the period, children earn for the household a wage, ω^i, that can differ per child depending on a set of characteristics, for example the gender of the child. The time spent in school leads to an increase in the "basic" wage in the second period. We denote by ρ the return on each unit of time spent in school, so that choosing t in period 1 gives as child income $(1-t)\omega^i$ in period 1, and a period 2 income of $(1+\rho t)\omega^i$.

The health risk arises in first period after a decision on the education of the child has been made. We model health risk as risk on wealth. With probability p, the parent falls ill in both periods (for the sake of simplicity – and consistent with the pattern of HIV/AIDS – we assume that no seropositivity is declared in the second period). The parent, if ill in the first period, loses an amount ε_1 of his revenue for that period and an amount ε_2 of his second period revenue.

The second period is modeled as a retirement period in which the parent does not personally earn money but rather relies on the appropriation of his child's income (in accordance with the literature on old-age security). We assume $\varepsilon_1 > \varepsilon_2$, that is, the monetary cost of illness is higher in the first period than in the second. This is because while in the second period the loss of revenue is only due to health costs, in the first period this loss is not only due to health costs but also due to the loss of work time which translates into an income loss.

[2] The model is easily extended to situations where the revenue of the parent comes from their own parents (in terms of educational investments). Results are robust to an overlapping-generation modeling with three periods (childhood, working age and retirement) in which the child is born in good health and can only fall ill in the second period.

The key mechanism of our model is the assumption that the proportion of child income appropriated by the parent in the second period differs according to the parent's health status and that this proportion also differs across children. The variation may be explained for example, by gender (see Maccini and Yang 2009 or Duflo 2000). To be precise, we assume that the parent appropriates a proportion δ of his child's revenue if healthy and a proportion $\delta + \Delta^i$ if unhealthy, where Δ^i is not necessarily homogeneous.

If we set β as the discount factor between the two periods, the problem of the parent in period 1 becomes:

$$\begin{aligned} \max_t \quad & p\left[u\left(\underline{C_1}\right) + \beta u\left(\underline{C_2}\right)\right] + (1-p)\left[u\left(\overline{C_1}\right) + \beta u\left(\overline{C_2}\right)\right] \\ \text{where} \quad & \overline{C_1} = (1-t)\omega^i + \omega \\ & \underline{C_1} = (1-t)\omega^i + \omega - \varepsilon_1 \\ & \overline{C_2} = \delta(1+\rho t)\omega^i \\ & \underline{C_2} = (\delta + \Delta^i)(1+\rho t)\omega^i - \varepsilon_2 \end{aligned} \tag{1}$$

For the model to remain tractable and realistic, we further assume that, if healthy, the parent is always richer in the first period than in the second period, that is $\overline{C_1} > \overline{C_2}$ (which corresponds to $\omega > \delta(1+\rho)\omega^i$) and always better off healthy than unhealthy, that is $\overline{C_1} > \underline{C_1}$ and $\overline{C_2} > \underline{C_2}$ (which is written as $\Delta^i(1+\rho)\omega^i < \varepsilon_2$).

To capture the role of "changing appropriation", we first analyze the baseline model where appropriation is the same irrespective of the health status of the parent, i.e. $\Delta^i = 0$. We compare this to the case where, through appropriation, education can play an insurance role $\Delta^i > 0$. We show that while optimal education is decreasing with disease prevalence in the baseline case, the asymmetric feature of appropriation can lead to a positive relationship.

2.1 The benchmark case: no insurance role for education ($\Delta^i = 0$)

Let us first assume that appropriation of child income by the parent does not increase in case of disease, that is $\Delta^i = 0$. In such a case, the first order condition of (1) writes:

$$-p\omega_i u'\left(\underline{C_1}\right) + \beta p \delta \rho \omega^i u'\left(\underline{C_2}\right) - (1-p)\omega_i u'\left(\overline{C_1}\right) + \beta(1-p)\delta\rho\omega^i u'\left(\overline{C_2}\right) = 0$$

That is:

$$\frac{\mathbb{E}(u'(C_1))}{\mathbb{E}(u'(C_2))} = \beta\delta\rho$$

To analyze the effect of disease prevalence, p, on the optimal time spent in school, t, we define

$$f(t,p) \equiv \frac{\mathbb{E}(u'(C_1))}{\mathbb{E}(u'(C_2))} - \beta\delta\rho$$

and use the Implicit Function Theorem:

$$\frac{\partial t}{\partial p} = -\frac{\partial f/\partial p}{\partial f/\partial t}$$

to obtain,

$$\begin{aligned}
\frac{\partial f}{\partial p} &= \frac{\left(u'\left(\underline{C_1}\right) - u'\left(\overline{C_1}\right)\right)\mathbb{E}(u'(C_2)) - \left(u'\left(\underline{C_2}\right) - u'\left(\overline{C_2}\right)\right)\mathbb{E}(u'(C_1))}{\left[\mathbb{E}(u'(C_2))\right]^2} \\
&= \frac{u'\left(\underline{C_1}\right)u'\left(\overline{C_2}\right) - u'\left(\overline{C_1}\right)u'\left(\underline{C_2}\right)}{\left[\mathbb{E}(u'(C_2))\right]^2} \qquad (2)
\end{aligned}$$

When $\Delta^i = 0$, $\underline{C_2} = \delta(1+\rho t)\omega^i - \varepsilon_2$ and $\overline{C_2} - \underline{C_2} = \epsilon_2$ for any t. As $\overline{C_1} - \underline{C_1} = \epsilon_1 \; \forall t$, the income risk faced by the parent is independent of the level of education he chooses for his child. Therefore, the educational choice reduces to a wealth transfer between the two periods. To abstract from the second order effect of wealth on risk aversion, we consider Constant Absolute Risk Aversion (CARA) preferences:

$$u(C) = -\frac{1}{\alpha}e^{-\alpha C}$$

It then turns out that:

Proposition 1. *If the proportion of child revenue appropriated by the parent does not increase when ill, i.e. $\Delta^i = 0$, the level of education, t, decreases with disease prevalence, p.*

Proof. See Appendix A. □

This theoretical result is as a direct consequence of the income effects and intertemporal smoothing decisions obtained in this framework. A health shock implies, first of all, household impoverishment, with income losses of ϵ_1 and ϵ_2. For $\epsilon_1 > \epsilon_2$, we have that the education decision is downsized in the face of diseases, to compensate potential income-loss in period 1. The slope between health risk, p, and the level of education, t, is then negative as in Fortson (2011), who uses a different micro modeling technique (human capital decision and mortality risk).

2.2 Assuming an insurance role for education ($\Delta^i > 0$)

If we now assume that the part of child revenue appropriated in period 2 by the parent is higher in case of ill health, i.e. $\Delta^i > 0$, the first order condition of program (1) becomes:

$$g(t,p) \equiv \frac{\mathbb{E}(u'(C_1)) - p\beta\Delta^i\rho u'(\underline{C_2})}{\mathbb{E}(u'(C_2))} - \beta\delta\rho = 0$$

with,

$$\begin{aligned}
\frac{\partial g}{\partial p} &= \frac{\partial f}{\partial p} - \frac{\beta\Delta^i\rho u'(\underline{C_2})\mathbb{E}(u'(C_2)) - p\beta\Delta^i\rho u'(\underline{C_2})(u'(\underline{C_2}) - u'(\overline{C_2}))}{[\mathbb{E}(u'(C_2))]^2} \\
&= \frac{u'\left(\underline{C_1}\right)u'\left(\overline{C_2}\right) - u'\left(\overline{C_1}\right)u'\left(\underline{C_2}\right) - \beta\Delta^i\rho u'(\underline{C_2})u'(\overline{C_2})}{[\mathbb{E}(u'(C_2))]^2}
\end{aligned} \tag{3}$$

Note here that two forces are at play when we compare (3) and (2): (i) an extra negative term appears in the numerator of (3) and (ii) Δ^i (positive) increases $\underline{C_2}$ so that we now have $\underline{C_2} > \overline{C_2} - \epsilon_2$. On the one hand, education has a higher return for the parent due to their ability to appropriate more, but on the other hand, the extra (expected) wealth in the second period reduces the need for education for intertemporal consumption smoothing.

Moreover

$$\frac{\partial g}{\partial t} = \frac{\partial f}{\partial t} - \frac{p\beta\Delta^i\rho(\delta+\Delta^i)\rho\omega^i u''(\underline{C_2})\mathbb{E}(u'(C_2)) - p\beta\Delta^i\rho u'(\underline{C_2})\left[p(\delta+\Delta^i)\rho\omega^i u''(\underline{C_2}) + (1-p)\delta\rho\omega^i u''(\overline{C_2})\right]}{[\mathbb{E}(u'(C_2))]^2}$$

Assuming a CARA utility function this simplifies to:

$$\frac{\partial g}{\partial t} = \frac{\partial f}{\partial t}_{|\Delta^i=0} - \frac{p\Delta^i\rho\omega^i}{[\mathbb{E}(u'(C_2))]^2}\left\{u''\left(\underline{C_2}\right)\mathbb{E}\left(u'\left(C_1\right)\right) + (1-p)\beta\rho\Delta^i u''\left(\underline{C_2}\right)u'(\overline{C_2})\right\} > 0$$

We therefore end up with,

$$\frac{\partial t}{\partial p} = \frac{1}{\omega^i}\frac{u'\left(\underline{C_1}\right)u'\left(\overline{C_2}\right) - u'\left(\underline{C_2}\right)\left[u'\left(\overline{C_1}\right) + \beta\Delta^i\rho u'(\overline{C_2})\right]}{\mathbb{E}(u''(C_1))\mathbb{E}(u'(C_2)) + \delta\rho\mathbb{E}(u''(C_2))\mathbb{E}(u'(C_1)) + p\Delta^i\rho\omega^i\left[u''\left(\underline{C_2}\right)\mathbb{E}(u'(C_1)) + (1-p)\beta\rho\Delta^i u''(\underline{C_2})u'(\overline{C_2})\right]} \tag{4}$$

that leads to the following proposition

Proposition 2. *When the proportion of child revenue appropriated by the parent is higher in the case of ill health ($\Delta^i > 0$), the level of education decreases with disease prevalence (in the case of CARA preferences) if and only if agents are impatient enough, that is if and only if*

$$\beta < \frac{1}{\Delta^i\rho}\left[\frac{u'(\underline{C_1})}{u'(\underline{C_2})} - \frac{u'(\overline{C_1})}{u'(\overline{C_2})}\right] \equiv \overline{\beta}$$

In the other case, that is if $\beta > \overline{\beta}$ (with $\overline{\beta} > 0$ whenever $\epsilon_1 > \epsilon_2$), the level of education increases in p.

Remark 1. *The role of changing appropriation (that is the "insurance role" of education) can be inferred from this condition. When Δ^i is positive, the negative slope between prevalence and education (found when $\Delta^i = 0$, see Proposition 1) is more difficult to obtain. For high level of β ($\beta > \overline{\beta}$), we even find that the sign changes, with the slope becoming positive.*

The intuition behind this result is quite simple. Assuming a change in the degree of appropriateness under bad health gives education an insurance role, calling for an increase in education when risk increases. The total effect of an increase in disease prevalence depends on whether this "insurance effect" dominates the consumption smoothing effect found in the previous section. According to proposition 2, the insurance effect dominates when the discount factor is large enough, that is, when the weight of the second period (the one in which education plays its insurance role) is large enough in the intertemporal utility function.

These theoretical results (Propositions 1 and 2) provides us with interesting testable predictions. Indeed, data would confirm these mechanisms if, after proxying health risk, we find that (i) the relationship between health risk and education is generally negative, but (ii) can become positive when education plays an important insurance role, for example for groups that are highly appropriable. In the next sections, we test these predictions on data assuming that child or parent/household characteristics can lead to different values of Δ^i, that is, to heterogeneity in the insurance role of education.

3 Data

We use the nationally representative cross-sectional data set, the Demographic and Health Survey (DHS), for 17 Sub-Saharan African countries, namely Burkina Faso (2003), Cameroon (2004), Congo Kinshasa (2007), Ethiopia (2005), Ghana (2003), Guinea (2005), Kenya (2008/09), Lesotho (2009), Liberia (2007), Malawi (2010), Mali (2006), Niger (2006), Senegal (2005), Sierra Leone (2008), Swaziland (2006/07), Zambia (2007) and Zimbabwe (2005/06). In this survey, data is collected at both the individual and household level. The DHS has been conducted in developing countries since 1984 with the aim of providing countries with data needed to monitor and evaluate population, health and nutrition programs on a regular basis. The data is collected usually every 5 years though a few are collected over shorter time intervals. It contains household data on basic characteristics of members of each household and also specific data on both male and female household members between the ages 15 and 49 inclusive. These surveys include testing for the HIV/AIDS status of members of the household 15 to 49 years of age. The sampling frame used in most DHS is, by definition, a list of non-overlapping area units with a majority of DHS sample designs clustered and stratified.

We test for the impact of HIV prevalence at the community level on educational enrollment focusing on individuals between the ages of 6 and 22 inclusive. There are, for our analysis, 357873 individuals within 128575 households nested in 6814 communities which are in turn nested in 17 countries. We take a look at the relationship between the HIV prevalence rates (that proxies health risk) and enrollment rates at the community level, for each country of our sample. The HIV prevalence rates are calculated using individuals between the ages of 15 and 49 years inclusive. From Figure 1, there appears not to exist *a priori* a clear relationship between prevalence rates and enrollment rate. Depending on the country, this relationship can be positive, negative, or even insignificant.

Insert Figure 1 here

Our theoretical model suggests that differences in appropriation of child income can be the source of this observed heterogeneity. A first way to examine this mechanism is to look at the differences in the bi-dimensional relationship between HIV prevalence rates and enrollment rates when differentiating urban and rural areas (appropriation being easier in rural areas according to Nugent 1985). Figure 2 takes a first look at this relationship for Lesotho (one of the countries with the most striking figures[3]). In Lesotho, the relation between community prevalence and educational enrollment is either positive or negative depending on the type of place of residence. In figure 1 we have that the overall effect is slightly positive (certainly insignificant) but when we take type of place of residence into account, we find that this effect can be decomposed into a negative relationship in urban areas and a positive one in rural areas where appropriation of child income is much easier.

Insert Figure 2 here

This bi-dimensional analysis however needs to be confirmed by a deeper study, controlling for various confounding factors known to influence educational enrollment rate. Therefore, in the following, in addition to community prevalence rates and country enrollment rates, we control for child and household characteristics including

- **Child characteristics:** Age, gender, residency status and relation to head.
- **Household characteristics:** Proportion of children below 5, type of place of residence, income, gender of head, age of head, HIV status of the household and educational level of head.

Regarding these characteristics (see table 1), our sample contains children of average age 12.85 years old with an almost equal number of males and females. Most of these children are usual residents of their households with a little over half being children of the head of the household. Households have about 16.85% of their children below the ages of 5. Approximately 69.07% of households are located in rural areas with 74.14% of the households being headed by males. About 43.93% of the household heads are uneducated, while 5.9801% of the households have an HIV+ member.

Insert Table 1 here

Regarding educational level (primary, secondary or university) we have:

Insert Table 2 here

[3]See Appendix 3 for results for all countries

4 Empirical Model

We now analyze more precisely the effect of HIV risk, proxied by community level HIV prevalence, on education, i.e. the attendance status of children. Recall that our theoretical model predicts that this effect is generally negative but can turn out to be positive for households in which the part of child income appropriated increases in case of bad health (this intuition being that is it more likely among highly appropriable children). To study the impact of community HIV prevalence on the school attendance status of a child we use a multilevel model. The multilevel model allows us to take into account random components at the individual, household, community and country level. As our data is hierarchical in nature, ignoring the clustering will imply that the independence condition is violated. In such a case, the standard errors are underestimated (see Guo and Zhao (2000); results from the standard binary model can be found in Appendix D). The hierarchical structure in our data comes from the fact that we pool our data across countries and also from the design of the DHS, where individuals are nested within households which are in turn nested within communities. Due to the complex nature of our analysis, that is the existence of 4 levels, we do a preliminary test of the significance of the levels. We are able to include contextual HIV/AIDS factors at the community level. We observe y_{ijkl}, the enrollment status of individual i in household j in community k in country l where

$$y_{ijkl} = \begin{cases} 1 & \text{if currently enrolled or was enrolled in current year} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Let us denote by p_{ijlk} the probability that $y_{ijkl} = 1$ that is:

$$p_{ijkl} = Pr\left(y_{ijkl} = 1\right)$$

Then by denoting $\pi_{ijkl} = \ln\left(\frac{p_{ijkl}}{1-p_{ijkl}}\right)$ we have that

$$\pi_{ijkl} = \alpha_0 + \alpha X_{ijkl} + \beta Z_{jkl} + \gamma W_{kl} + \eta V_l + \theta_{ijkl} + \upsilon_{jkl} + \epsilon_{kl} + \phi_l \tag{6}$$

where α_0 is the intercept, α is the vector of coefficients for the covariates X_{ijkl}, β is the vector of coefficients associated with the covariates Z_{jkl}, γ the vector of coefficients for covariates W_{kl} and η the vector of coefficients for covariates V_l. The covariates, X_{ijkl} varies at the individual level, Z_{jkl} varies at the household level, W_{kl} at the community level and V_l at the country level. The error terms at the individual, household, community and country levels are given as θ_{ijkl}, υ_{jkl}, ϵ_{kl} and ϕ_l respectively and have means 0 and variances, σ^2_{in}, σ^2_{hh}, σ^2_{cm} and σ^2_{ct} respectively, (see Goldstein 1991). We first calculate the intra-class correlations to look at the contribution of each level to the overall variance. This variable is calculated from the empty model:

$$\pi_{ijkl} = \alpha_0 + \theta_{ijkl} + \upsilon_{jkl} + \epsilon_{kl} + \phi_l \tag{7}$$

Estimates of household, community and country level variances are used to calculate the intra-class correlation coefficients to determine the proportion of group-level variance compared to total variance, using the following formula:

$$\rho_{hh} = \frac{\sigma^2_{hh}}{\sigma^2_{hh} + \sigma^2_{cm} + \sigma^2_{ct} + \sigma^2_{in}} \tag{8}$$

$$\rho_{cm} = \frac{\sigma^2_{rg}}{\sigma^2_{hh} + \sigma^2_{cm} + \sigma^2_{ct} + \sigma^2_{in}} \tag{9}$$

and

$$\rho_{ct} = \frac{\sigma^2_{ct}}{\sigma^2_{hh} + \sigma^2_{cm} + \sigma^2_{ct} + \sigma^2_{in}} \tag{10}$$

where ρ_{hh}, ρ_{cm} and ρ_{ct} are the intra-class correlations at the household, community and country levels and σ^2_{in} is the variance at the individual level, where for the case of a multilevel logistic regression is fixed at $\frac{\pi^2}{3}$, i.e. 3.289 (see Hedeker and Gibbsons 1996).

The empty model gives us as variances for the four levels $\sigma^2_{in} = 3.2899$, $\sigma^2_{hh} = 0.7017$, $\sigma^2_{cm} = 0.9317$ and $\sigma^2_{ct} = 0.7487$. From equation (8) we obtain as intra-class correlation rates $\rho_{hh} = 12.3713$, $\rho_{cm} = 16.4263$ and $\rho_{ct} = 13.1999$ meaning that 12.3713% of the total variation is explained at the household level, with 16.4263% explained at the community level and 12.3713% at the country level. Based on this result, we include the four levels in our multilevel model. We now discuss the results we obtain from our multilevel model.

5 Results

Running our empirical model on Demographic and Health Surveys, we confirm previous findings (see table [?]). We indeed find that (i) having an HIV+ household member implies a negative impact on the probability of enrolling in school, a result that can be reconciled with Fortson (2011) (ii) living in rural areas also has a negative and significant effect on enrollment whereas (iii) being from a high income household has a positive and significant impact on the probability of being enrolled. Regarding, our variable of interest, we first find that, taking into account the whole sample, the probability of being enrolled in school is not significantly affected by the probability of adult household members becoming HIV+.

The results are nuanced if, as in Nugent (1985), and based on the second part of the theoretical model (allowing for heterogeneous appropriateness), we split our data based on gender and on type of place of residence of the household.

Insert Table 3 here

An increase in HIV prevalence (the probability of adult household members becoming HIV+) leads to a reduction in male enrollment but an increase in female children enrollment, though the effect for females are insignificant. According to our theoretical, this would come from the fact that females are more appropriable than their male counterparts. The result can be related to the findings of Maccini and Yang (2009) or De la Brière et al. (2002) who advocate that girls are often used as insurance assets (in the case of rainfall shocks for Maccini and Yang (2009)).

The same kind of results can be inferred from the split between urban and rural regions. As children are more easily appropriable in rural regions, we find that an increase in HIV-prevalence decreases the probability of enrollment in urban areas but increase it in rural areas. These two results on subsamples confirm the predictions of our theoretical model, as we find that for subgroup of children that are highly appropriable (female children and children in rural areas), the insurance role of education can lead to a positive (and sometimes significant) relationship between the probability of becoming HIV+ and the probability of being enrolled in school.

Insert Table 4 here

As a final check we split the data based on the age groups (see Table 4). We find that our key relationship (between HIV prevalence and educational enrollment) is positive and significant for children from 6 to 12, insignificant for children from 13 to 18 and negative and significant for children aged 19 to 22. This seems to confirm the insurance role plaid by children future income as the positive relationship highlighted in our theoretical model is mainly present for young children (who closely corresponds to the one modeled in the theoretical part).

The two last columns of Table 4 split data with respect to the degree of community HIV prevalence. We then find that our positive relationship is restricted to areas where HIV prevalence remains low. In our view, this underlines that the total effect consists of the insurance (positive) effect presented here and the usual negative effect (through household income effect and other macroeconomic mechanisms[4]); and that the positive effects can only dominate in low prevalence community.

6 Discussion and Conclusion

We have analyzed the impact of health risk, and specifically HIV/AIDS risk, on education taking into account the fact that education of children can act as a form of insurance against shocks, and in this particular case, health shocks in old-age. We first built a theoretical model where active adults invest in the education of their children. We find that in the absence of extra appropriation in case of ill health, an increase in the probability of falling ill leads to a reduction in the time children spend in school. However when there is an opportunity for extra appropriation in bad health events, we have that health risk may in some cases lead to an increase in the amount of education, especially when family's response in terms of required remittance (the extra amount appropriable) is very high.

We test the idea that an increase in the probability of falling ill can lead to an increase in educational investments and find that (i) for the overall population of 6 to 22-year-olds the effect is negative but that (ii) once the data is split between males and females the effect becomes positive for females. When we also split the data between rural and urban households we find a positive and significant effect in the rural areas.

We therefore verify the conventional result (a negative relationship between education and HIV prevalence, see Fortson (2011)) for a subsample of children: male children and children living in urban areas. But we enrich it for children of another type: female children and children in rural areas for whom the relationship seems to be positive (a result already emphasized by some: see

[4]Unicef 2000[26] underlines that the number of African children having lost their teachers because of AIDS was around 860 000 in 1999, It means that the replacement rate of teachers, due to the pandemic, is not large enough to provide a desirable level of education.

Bennell, Hyde and Swainson (2002) and Bennell (2004)). This empirical “puzzle”, i.e. this surprising non-monotonicity, opens the door to a new theoretical reasoning different from the usual one of consumption smoothing. Our interpretation of these results is that, depending on how appropriable the income of ones child is, education acts as an insurance asset, providing protection in bad events. This attribute could reverse the sign of the relationship between disease prevalence and schooling. Our theoretical model provides one possible explanation of such a mechanism.

These results suggest that the role of remittance systems would be an interesting avenue to explore. In our empirical setting, we have proxied differences in appropriateness by exogenous characteristics (gender, urban/rural areas) that have been proven in other papers to be related to appropriateness. However, more attention needs to be paid to our key mechanism of “changing appropriateness”, for example by analyzing to what extent (and based on which characteristics) remittances change depending on the health status of parents.

Last, regarding the policy implications of our work, it seems worth emphasizing that the general picture we obtain is that HIV/AIDS disease indeed has a negative impact on school enrollment. First, the fact that family members are infected (individual HIV status) is always associated with a decrease in the probability of schooling. Second, when we obtain a positive effect of HIV risk rate (community prevalence), we must keep in mind that the estimated coefficient evaluates the actual consolidation of different forces, sometimes positive (insurance effect) and generally negative (income effect, supply side effect on the educational sector). This is clearly demonstrated by the various stratifications of the estimations, showing that the positive effect only dominates for young children, in rural areas, in countries of low prevalence. In brief, the findings of this paper clearly do not support a decrease in schooling support in Least Developed Countries(LDC). To the contrary, the lesson is that supply-oriented policies should be favored, as the demand for education can be partly self-sustained, given the insurance effect.

References

[1] Becker, Gary S. and Robert J. Barro. 1988. “A Reformulation of the Economic Theory of Fertility.” The Quarterly Journal of Economics 103 (February): 1-25.

[2] Basu, Kaushik and Pham H. Van. 1998. “The Economics of Child Labor.” The American Economic Review 88 (June): 412-427.

[3] Bazen, Stephen and Claire Salmon. 2010. “The Impact of Parental Health on Child Labor: Evidence from Bangladesh.” Economic Bulletin 30 (April): 2549-2557.

[4] Bennell, Paul. 2004. “AIDS in Africa: Three Scenarios for the Education Sector.” Report Prepared for the UNAIDS Project, AIDS in Africa: Scenarios for the Future.

[5] Bennell, Paul, Karine Hyde and Nicola Swainson. 2002. “The Impact of HIV/AIDS Epidemic on the Education Sector in Sub-Saharan Africa: A Synthesis of the Findings and Recommendations of Three Country Studies.” Centre for International Education. University of Sussex Institute of Education.

[6] Conroy, Anne C., Malcolm J. Blackie, Alan Whiteside, Justin C. Malewezi and Jeffery D. Sachs. 2006. “Poverty, AIDS and Hunger: Breaking the Poverty Trap in Malawi.” 1st ed. New York: Palgrave MacMillan.

[7] De la Bri‘ere, Benedicte, Elisabeth Sadoulet, Alan de Janvry and Sylvie Lambert. 2002. “The roles of destination, gender, and household composition in explaining remittances: an analysis for the Dominican Sierra.” Journal of Development Economics 68 (August): 309 328.

[8] Duflo, Esther. 2000. “Child Health and Household Resources in South Africa: Evidence from the Old Age Pension Program.” The American Economic Review 90 (May): 393-398.

[9] Ehrlich, Isaac and Francis T. Lui. 1991. “Intergenerational Trade, Longevity, and Economic Growth.” Journal of Political Economy 95 (October): 1029-1059.

[10] Fafchamps, Marcel. 1992. “Solidarity Networks in Preindustrial Societies: Rational Peasants with a Moral Economy.” Economic Development and Cultural Change 41 (October): 147-174.

[11] Flores, Gabriela, Jaya Krishkumar, Owen O’Donnell and Evan van Doorslaer. 2008. “Coping with Healthcare Costs: Implications for the Measurement of Catastrophic Expenditures and Poverty.” Health Economics 17 (February): 1393-1412.

[12] Fortson, Jane. 2011. “Mortality Risk and Human Capital Investment: The Impact of HIV/AIDS in Sub-Saharan Africa.” Review of Economics and Statistics 93 (February): 1-15.

[13] Gertler, Paul and Jonathan Gruber. 2002. "Insuring Consumption against Illness." The American Economic Review. 92 (March): 51-70.

[14] Goldstein, Harvey. 1991. "Nonlinear Multilevel Models with an Application to Discrete Response Data." Biometrika 78 (March): 45-51.

[15] Guo, Guang and Hongxin Zhao. 2000. "Multilevel Modeling for Binary Data." Annual Review of Sociology 26 (August): 441-462.

[16] Harrower, Sarah and John Hoddinott. 2005. "Consumption Smoothing in the Zone Lacustre, Mali." Journal of African Economies 14 (December): 489-519.

[17] Hedeker, Donald and Robert D. Gibbsons. 1996. "MIXOR: A Computer Programme for Mixed Effects Ordinal Regression Analysis." Computer Methods and Programs in Biometrics 49 (March): 157-176.

[18] Kawabata, Kei, Ke Xu and Guy Carrin. 2002. "Preventing Impoverishment through Protection against Catastrophic Health Expenditure." Bulletin World Health Organisation 80: 612 - 630.

[19] Maccini, Sharon and Dean Yang. 2009. "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall." The American Economic Review 99 (June): 1006-1026.

[20] McIntyre, Di, Michael Thiede, Göran Dahlgren and Margret Whitehead. 2006. "What are the Economic Consequences for Households of Illness and of Paying for Healthcare in Low- and middle-income Country Contexts?" Social Science and Medicine 62 (February): 858-865.

[21] Michel, Philippe and Bertrand Wigniolle. 2007. "On Efficient Child Making." Economic Theory 31 (May): 307-326.

[22] Nugent, Jeffrey B. 1985. "The Old Age Security Motive for Fertility." Population and Development Review 11 (March): 75-97.

[23] Park, Cheolsung. 2006. "Risk Pooling between Households and Risk Coping Measure in Developing Countries: Evidence from Rural Bangladesh." Economic Development and Cultural Change 54 (January): 423-457.

[24] Rapoport, Hillel and Frederic Docquier. 2006. "The Economics of Migrants' Remittances, on the Economics of Giving, Reciprocity and Altruism." Handbook of Economics of Giving, Altruism and Reciprocity 2, 1st ed.

[25] Schoonbroodt, Alice and Michele Tertilt. 2010. "Who Owns Children and Does it Matter." NBER Working Paper 15663.

[26] United Nations Children's Fund (UNICEF). 2000. "The Progress of Nations," New York.

[27] Wagstaff, Adam and Eddy van Doorslaer. 2003. "Catastrophe and Impoverishment in Paying for Health Care: With Applications to Vietnam 1993-1998." Health Economics 12 (November): 921-934.

[28] Xu, Ke, David B. Evans, Kei Kawabata, Riadh Zeramdini, Jan Klavus and Christopher J. L. Murray. 2003. "Household Catastrophic Health Expenditure: A Multicountry Analysis." Lancet 362 (July): 111-117.

A Proof of Proposition 1

We have:

$$\begin{aligned}
\frac{\partial f}{\partial p} &= \frac{\left(u'\left(\underline{C_1}\right) - u'\left(\overline{C_1}\right)\right)\mathbb{E}(u'(C_2)) - \left(u'\left(\underline{C_2}\right) - u'\left(\overline{C_2}\right)\right)\mathbb{E}(u'(C_1))}{\left[\mathbb{E}(u'(C_2))\right]^2} \\
&= \frac{u'\left(\underline{C_1}\right)u'\left(\overline{C_2}\right) - u'\left(\overline{C_1}\right)u'\left(\underline{C_2}\right)}{\left[\mathbb{E}(u'(C_2))\right]^2}
\end{aligned} \tag{11}$$

Therefore, $\frac{\partial f}{\partial p}$ is of the sign of $\frac{u'\left(\underline{C_1}\right)}{u'\left(\overline{C_1}\right)} - \frac{u'\left(\underline{C_2}\right)}{u'\left(\overline{C_2}\right)}$, that is, in the case of $\Delta_i = 0$, the sign of $\frac{u'\left(\overline{C_1}-\varepsilon_1\right)}{u'\left(\overline{C_1}\right)} - \frac{u'\left(\overline{C_2}-\varepsilon_2\right)}{u'\left(\overline{C_2}\right)}$

Noticing that $\frac{u'(C-\varepsilon)}{u'(C)}$ is constant in C in the case of Constant Absolute Risk Aversion, it turn out that: $\frac{u'\left(\overline{C_1}-\varepsilon_1\right)}{u'\left(\overline{C_1}\right)} - \frac{u'\left(\overline{C_2}-\varepsilon_2\right)}{u'\left(\overline{C_2}\right)} \geq 0$ as $\varepsilon_1 > \varepsilon_2$.

Lemma 1. *If $u(.)$ exhibits CARA, $\frac{\partial f}{\partial p}_{|\Delta_i=0} \geq 0$*

Note that the previous lemma also hold for preferences exhibiting Increasing Absolute Risk Aversion and that the result is ambiguous for Decreasing absolute risk aversion

As, moreover,

$$\frac{\partial f}{\partial t}_{|\Delta_i=0} = \frac{-\omega^i\mathbb{E}(u''(C_1))\mathbb{E}(u'(C_2)) - \delta\rho\omega^i\mathbb{E}(u''(C_2))\mathbb{E}(u'(C_1))}{\left[\mathbb{E}(u'(C_2))\right]^2} > 0$$

proposition 1 holds.

B Description of Variables and Summary Statistics

Insert Table 5 here

Insert Table 6 here

Insert Table 7 here

C Community Enrollment vs Prevalence Over Type of Place of Residence for All Countries

Insert Figure 3 here

D Results from the Binary Model

Insert Table 8 here

E Graphs and Figures

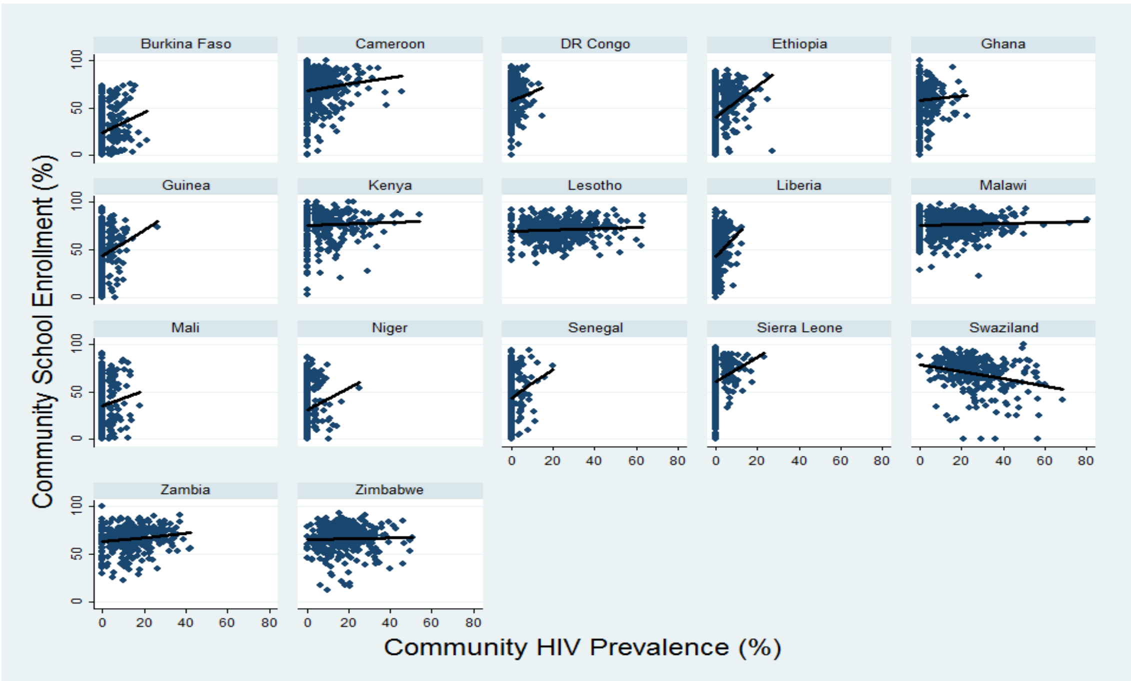

Figure 1: Community Enrollment vs Prevalence Per Country

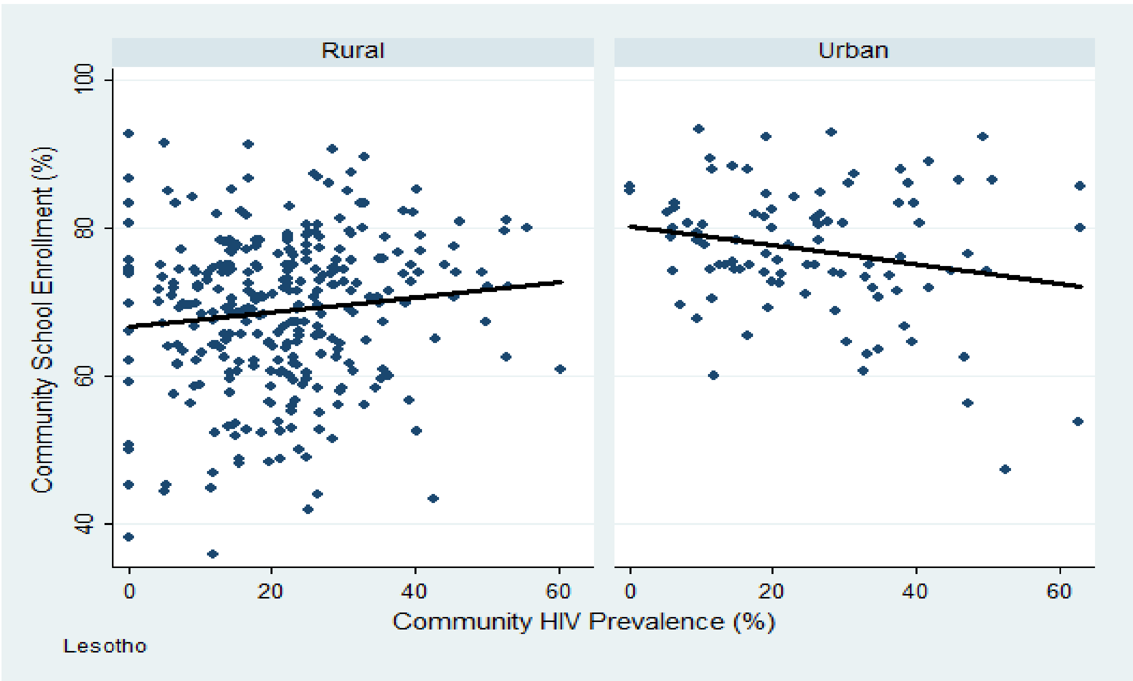

Figure 2: Community Enrollment vs Prevalence Over Type of Place of Residence for Lesotho

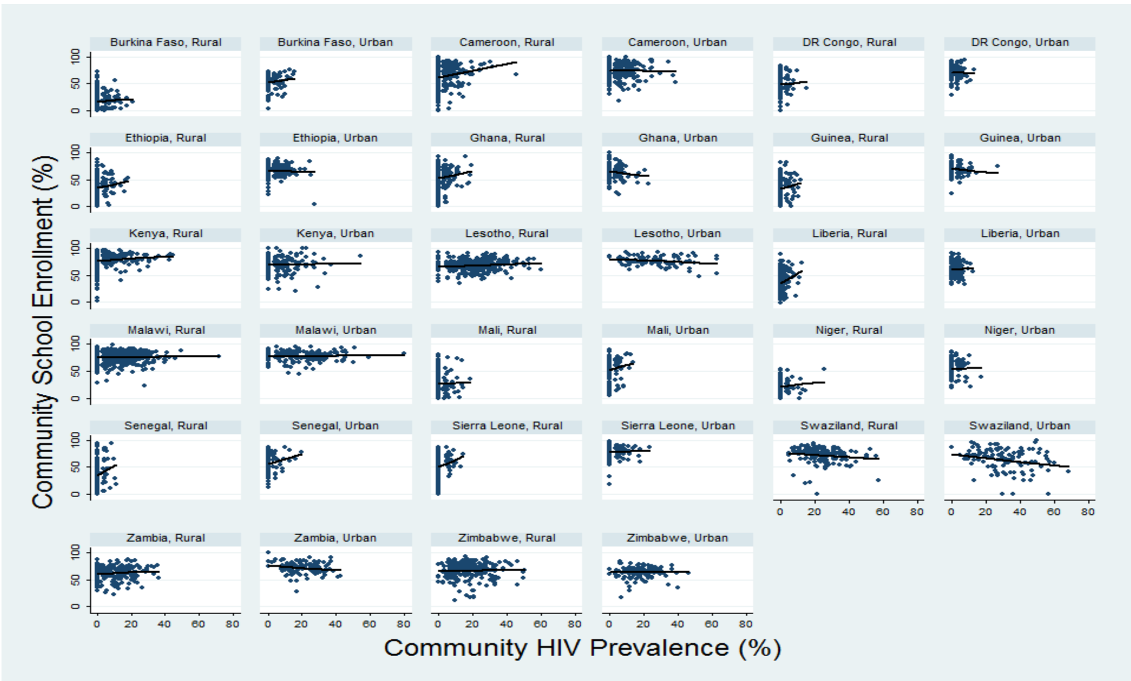

Figure 3: Community Enrollment vs Prevalence Over Type of Place of Residence for All Countries

E.1 Tables

Variable	N	Mean	Std Dev
Age	363607	12.8557	4.7883
Male	363592	0.4979	0.5000
Usual Resident	363351	0.9746	0.1574
Child of Head	363532	0.6256	0.4840
Prop. of Children < 5	130165	0.1685	0.1560
Rural	130165	0.6907	0.4622
Male Head	130165	0.7414	0.4379
Head 40 - 59 years	130165	0.3968	0.4892
Head 60 - 79 years	130165	0.1778	0.3823
Head $\geq$ 80 years	130165	0.0205	0.1416
Middle Income	130165	0.2004	0.4003
High Income	130165	0.4125	0.4923
Household HIV+	130165	0.0598	0.2371
Household HIV Status Unknown	130165	0.5149	0.4998
Head Uneducated	129081	0.4393	0.4963

Table 1: Summary Statistics

	No Education	Primary	Secondary	Higher	Total
6 to 12	38.79	60.78	0.43	0.00	100.00
13 to 18	23.57	53.17	23.18	0.08	100.00
19 to 22	28.66	32.40	36.66	2.27	100.00
Total	32.18	53.70	13.72	0.39	100.00

Table 2: Highest Educational Level of Individuals Per Age Group

Fixed Effects					
Variable	**Total**	**Male**	**Female**	**Urban**	**Rural**
Intercept	-3.3371^{***}	-2.4997^{***}	-3.6439^{***}	-1.7544^{***}	-4.8271^{***}
	(0.1397)	(0.0853)	(0.2626)	(0.3045)	(0.1768)
Age	1.1302^{***}	1.1362^{***}	1.1719^{***}	0.9639^{***}	1.2402^{***}
	(0.0069)	(0.0094)	(0.0105)	(0.0119)	(0.0086)
Age-Squared	-0.0480^{***}	-0.0464^{***}	-0.0517^{***}	-0.0410^{***}	-0.0528^{***}
	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.0003)
Male	0.4530^{***}			0.5300^{***}	0.4173^{***}
Usual Resident	0.1763^{***}	0.2892^{***}	0.0493	0.2249^{***}	0.0916^{**}
	(0.0330)	(0.0486)	(0.0458)	(0.0582)	(0.0404)
Child of Head	0.6871^{***}	0.4069^{***}	0.8397^{***}	0.8605^{***}	0.5921^{***}
	(0.0120)	(0.0166)	(0.0169)	(0.0206)	(0.0147)
Household HIV+	-0.0487^{*}	-0.0683^{*}	-0.0281	0.0122	-0.0877^{***}
	(0.0274)	(0.0362)	(0.0368)	(0.0467)	(0.0337)
Household HIV Unknown	-0.0266^{**}	-0.0063	-0.0388^{**}	-0.0150	-0.0337^{**}
	(0.0129)	(0.0164)	(0.0174)	(0.0244)	(0.0152)
Rural	-0.8149^{***}	-0.8188^{***}	-0.8352^{***}		
	(0.0312)	(0.0336)	(0.0363)		
Middle Income	0.3195^{***}	0.3155^{***}	0.3739^{***}	0.2657^{***}	0.3616^{***}
	(0.0171)	(0.0217)	(0.0229)	(0.0565)	(0.0180)
High Income	0.7941^{***}	0.8002^{***}	0.8855^{***}	0.9762^{***}	0.7527^{***}
	(0.0192)	(0.0300)	(0.0257)	(0.0522)	(0.0211)
Prop. of Child. < 5	-0.8896^{***}	-0.0335	-1.3894^{***}	-1.5303^{***}	-0.6006^{***}
	(0.0462)	(0.0629)	(0.0624)	(0.0876)	(0.0543)
Head Uneducated	-0.7298^{***}	-0.7720^{***}	-0.7910^{***}	-0.9386^{***}	-0.6136^{***}
	(0.0158)	(0.0201)	(0.0210)	(0.0289)	(0.0188)
Head 40-59	0.1765^{***}	0.1341^{***}	0.2147^{***}	0.2731^{***}	0.1235^{***}
	(0.0144)	(0.0190)	(0.0195)	(0.0262)	(0.0172)
Head 60-79	0.2958^{***}	0.2051^{***}	0.3758^{***}	0.3917^{***}	0.2308^{***}
	(0.0184)	(0.0239)	(0.0249)	(0.0352)	(0.0217)
Head > 80	0.4318^{***}	0.3225^{***}	0.5504^{***}	0.6588^{***}	0.3078^{***}
	(0.0426)	(0.0551)	(0.0568)	(0.0916)	(0.0482)
Head Male	-0.2703^{***}	-0.2368^{***}	-0.3440^{***}	-0.2600^{***}	-0.2552^{***}
	(0.0146)	(0.0194)	(0.0195)	(0.0258)	(0.0172)
Cluster HIV Prevalence	0.0009	-0.0035^{**}	0.0017	-0.0051^{**}	0.0108^{***}
	(0.0017)	(0.0016)	(0.0021)	(0.0024)	(0.0023)
Country Enrollment Rate	0.0624^{***}	0.0570^{***}	0.0693^{***}	0.0271^{***}	0.0758^{***}
	(0.0023)	(0.0010)	(0.0044)	(0.00582)	(0.0029)
Random Effects					
Household					
Int. Variance	1.0133	0.9574	0.9006	1.1560	0.9283
Int. Std. Deviation	1.0066	0.9784	0.9490	1.0752	0.9635
Community					
Int. Variance	0.8990	0.8651	1.0008	0.3874	0.9538
Int. Std. Deviation	0.9481	0.9301	1.0004	0.6224	0.9766
Country					
Int. Variance	0.0173	-	0.0744	0.0963	0.0290
Int. Std. Deviation	0.1317	-	0.2728	0.3103	0.1703
Number of Obs.	357873	178561	179332	105563	252310
Groups: Household	128575	93002	101415	37060	91515
Groups: Cluster	6814	6812	6810	2291	4523
Groups: Country	17	17	17	17	17

Table 3: Impact of HIV Prevalence

Fixed Effects					
Variable	**6 - 12**	**13 - 18**	**19 - 22**	**Low Prevalence**	**High Prevalence**
Intercept	-3.4162^{***}	-3.2379^{***}	-2.5942^{***}	-3.1362^{***}	-2.9248^{***}
	(0.6458)	(0.3262)	(0.6636)	(0.1321)	(0.2992)
Age	2.9692^{***}	0.7783^{***}	-2.5452^{***}	1.0722^{***}	1.2356^{***}
	(0.0391)	(0.1019)	(0.5104)	(0.0080)	(0.0140)
Age-Squared	-0.1406^{***}	-0.0395^{***}	0.0520^{***}	-0.0445^{***}	-0.0552^{***}
	(0.0022)	(0.0033)	(0.0125)	(0.0003)	(0.0005)
Male	0.1929^{***}	0.6193^{***}	0.8772^{***}	0.5416^{***}	0.1860^{***}
	(0.0152)	(0.0170)	(0.0265)	(0.0114)	(0.0202)
Usual Resident	0.1444^{*}	0.1801^{***}	-0.3535^{***}	0.1085^{**}	0.0856^{*}
	(0.0767)	(0.0491)	(0.0561)	(0.0476)	(0.0486)
Child of Head	0.4001^{***}	0.8078^{***}	0.6162^{***}	0.6996^{***}	0.7241^{***}
	(0.0204)	(0.0198)	(0.0313)	(0.0142)	(0.0233)
Household HIV+	-0.0501	-0.0806^{*}	-0.2088^{***}	-0.0176	-0.0900^{***}
	(0.0468)	(0.0413)	(0.0556)	(0.0599)	(0.0340)
Household HIV Unknown	-0.0410^{**}	-0.1021^{***}	-0.0088	-0.0211	-0.0594^{**}
	(0.0193)	(0.0198)	(0.0288)	(0.0147)	(0.0285)
Rural	-1.1558^{***}	-0.7229^{***}	-0.6304^{***}	-1.0414^{***}	-0.2272^{***}
	(0.0450)	(0.0381)	(0.0423)	(0.0392)	(0.0498)
Middle Income	0.4307^{***}	0.4437^{***}	0.1872^{***}	0.3289^{***}	0.3351^{***}
	(0.0248)	(0.0260)	(0.0421)	(0.0200)	(0.0340)
High Income	1.0130^{***}	0.8784^{***}	0.8174^{***}	0.7943^{***}	0.8823^{***}
	(0.0287)	(0.0287)	(0.0410)	(0.0226)	(0.0378)
Prop. of Child. <5	-0.0785	-1.0899^{***}	-2.6911^{***}	-0.6944^{***}	-1.4031^{***}
	(0.0710)	(0.0753)	(0.1069)	(0.0546)	(0.0907)
Head Uneducated	-0.9036^{***}	-0.8728^{***}	-0.8385^{***}	-0.7500^{***}	-0.6491^{***}
	(0.0236)	(0.0236)	(0.0350)	(0.0185)	(0.0315)
Head 40-59	0.0014	0.2395^{***}	0.6403^{***}	0.1878^{***}	0.1811^{***}
	(0.0214)	(0.0242)	(0.0375)	(0.0171)	(0.0282)
Head 60-79	0.0676^{**}	0.3239^{***}	0.6862^{***}	0.2537^{***}	0.4100^{***}
	(0.0289)	(0.0285)	(0.0432)	(0.0219)	(0.0362)
Head >80	0.2274^{***}	0.5024^{***}	0.6635^{***}	0.4326^{***}	0.4312^{***}
	(0.0670)	(0.0610)	(0.0958)	(0.0515)	(0.0790)
Head Male	-0.2526^{***}	-0.2967^{***}	-0.2743^{***}	-0.2929^{***}	-0.2225^{***}
	(0.0235)	(0.0219)	(0.0302)	(0.0182)	(0.0259)
Cluster HIV Prevalence	0.0121^{***}	-0.0031	-0.0065^{***}	0.0380^{***}	-0.0025
	(0.0027)	(0.0021)	(0.0023)	(0.0083)	(0.0022)
Country Enrollment Rate	0.0892^{***}	0.0635^{***}	0.0186^{*}	0.0610^{***}	0.0567^{***}
	(0.0109)	(0.0055)	(0.0112)	(0.0021)	(0.0048)
Random Effects					
Household					
Int. Variance	1.52714	0.7429	0.4828	1.0164	1.1390
Int. Std. Deviation	1.2358	0.8619	0.6948	1.0082	1.0673
Community					
Int. Variance	1.4231	0.9406	0.6329	0.9899	0.5351
Int. Std. Deviation	1.1929	0.9698	0.7956	0.9950	0.7315
Country					
Int. Variance	0.4925	0.1188	0.5212	0.0113	0.0765
Int. Std. Deviation	0.7018	0.3447	0.7219	0.1064	0.2766
Number of Obs.	184164	116685	57024	255785	102088
Groups: Household	94988	73415	46124	88929	39646
Groups: Cluster	6802	6802	6773	4607	2207
Groups: Country	17	17	17	17	17

Table 4: Impact of HIV Prevalence (Age and Prevalence)

Variable	**Measure**
	Outcome Variable
Currently Enrolled	Coded as 1 if individual is currently enrolled or was enrolled in current year, 0 otherwise
	Explanatory Variables
Age	Age of Child centered at grand mean
Age^2	The squared age of the child centered at grand mean
Usual Resident	Coded as 1 if the individual is a usual resident of the household and 0 if not
Child of Head	Coded as 1 if the individual is the child of the household head and 0 otherwise
Prop. of Children < 5	Proportion of household members who are less than 5 years old
Rural	Coded as 1 if the household is located in a rural area and 0 otherwise
Male Head	Coded as 1 if the household is headed by a man and 0 otherwise
Head 40 - 59 years	Coded as 1 if the household head is between the ages of 40 and 59 inclusive and 0 otherwise
Head 60 - 79 years	Coded as 1 if the household head is between the ages of 60 and 79 inclusive and 0 otherwise
Head $\geq$ 80 years	Coded as 1 if the household head is at least 80 years and 0 otherwise[5]
Middle Income	Coded as 1 if the household is middle income and 0 otherwise
High Income	Coded as 1 if the household is high income and 0 otherwise[6]
Head HIV+	Coded as 1 if household head is HIV+ and 0 if negative or unknown
Head HIV status unknown	Coded as 1 if household head's HIV status is unknown and 0 if head is either HIV+ or HIV-
Head Uneducated	Coded as 1 if household head is uneducated and 0 otherwise
	Cluster Level Contextual Variable
Cluster HIV Rate	The weighted HIV prevalence rates at the community level
	Country Level Contextual Variable
Country Enrollment Rate	The weighted Enrollment rates at the national level

Table 5: Variable Description

Community Enrollment	**Burkina Faso**	**Cameroon**	**D.R. Congo**	**Ethiopia**	**Ghana**	**Guinea**	**Kenya**	**Lesotho**	**Liberia**	**Malawi**	**Mali**	**Niger**	**Senegal**	**Sierra Leone**	**Swaziland**	**Zambia**	**Zimbabwe**
Mean	25.4629	69.9848	58.8845	43.7872	58.5878	45.7571	76.0471	70.7228	46.8559	76.2974	36.5628	32.3360	44.7868	62.6147	68.7392	66.5060	66.6448
Std. Dev.	20.5263	19.4949	17.8182	22.2725	18.1275	23.5128	15.7913	10.4968	20.6388	9.6688	22.7077	20.6259	23.2337	21.6637	17.2218	12.8698	12.8975
Median	20.2254	75	60.0585	44.7761	62.1820	46.1538	80	72.1576	48.7955	77.4194	35.1648	30.0794	46.0044	67.6471	72.2222	68.5185	68.1818
Obs	400	466	300	534	412	292	395	400	298	847	407	342	376	353	275	319	398

Table 6: Summary Statistics of Cluster Enrollment Rate (%)

Community Enrollment	**Burkina Faso**	**Cameroon**	**D.R. Congo**	**Ethiopia**	**Ghana**	**Guinea**	**Kenya**	**Lesotho**	**Liberia**	**Malawi**	**Mali**	**Niger**	**Senegal**	**Sierra Leone**	**Swaziland**	**Zambia**	**Zimbabwe**
Mean	1.8167	5.7832	1.3277	2.0049	2.2586	1.5534	7.0473	22.2591	1.6783	10.7839	1.2155	1.0425	0.8874	1.5185	26.8747	14.2148	18.0582
Std. Dev.	3.7134	7.0585	2.4655	4.1870	3.8330	3.2699	9.4101	12.6784	2.4476	11.6696	2.8586	2.8892	2.5364	3.5170	12.2443	10.2653	9.3527
Median	0	4.2834	0	0	0	0	4.6556	21.7038	0	7.84921	0	0	0	0	25.5852	13.5156	17.0611
Obs	400	466	300	535	412	292	395	400	298	847	407	342	376	353	275	319	398

Table 7: Summary Statistics of Cluster Prevalence Rate (%)

Variable	Total	Male	Female	Urban	Rural	6 - 12	13 - 18	19 - 22	Low Prevalence	High Prevalence
Intercept	−2.0635***	−1.6603***	−2.1554***	−1.2963***	−3.0074***	−1.8746***	−1.9405***	−1.6266***	−1.9407***	−2.0589***
	(0.0377)	(0.0541)	(0.0529)	(0.0703)	(0.0452)	(0.0689)	(0.0605)	(0.0827)	(0.0489)	(0.0738)
Age	0.8894***	0.8771***	0.9306***	0.7618***	0.9759***	1.9408***	0.7280***	−1.7640***	0.8297***	0.9867***
	(0.0057)	(0.0078)	(0.0086)	(0.0100)	(0.0072)	(0.0303)	(0.0867)	(0.4427)	(0.0066)	(0.0117)
Age-Squared	−0.0378***	−0.0358***	−0.0409***	−0.0323***	−0.0415***	−0.0917***	−0.0343***	0.0349**	−0.0344***	−0.0439***
	(0.0002)	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0017)	(0.0028)	(0.0108)	(0.0003)	(0.0004)
Male	0.3364***			0.4068***	0.3113***	0.1102***	0.4715***	0.7057***	0.3968***	0.1457***
	(0.0082)			(0.0152)	(0.0099)	(0.0117)	(0.0142)	(0.0229)	(0.0095)	(0.0170)
Usual Resident	0.0132	0.2271***	−0.1859***	0.0722	−0.0142	−0.1467*	0.1296**	−0.2854***	−0.0318	−0.0497
	(0.0265)	(0.0390)	(0.0366)	(0.0471)	(0.0322)	(0.0570)	(0.0400)	(0.0457)	(0.0385)	(0.0383)
Child of Head	0.4289***	0.2549***	0.5545***	0.5930***	0.3592***	0.1661***	0.5715***	0.4751***	0.4332***	0.5008***
	(0.0093)	(0.0131)	(0.0134)	(0.0162)	(0.0114)	(0.0145)	(0.0160)	(0.0266)	(0.0109)	(0.0185)
Household HIV+	−0.0380	−0.0586*	−0.0265	−0.0044	−0.0558*	0.0132	−0.0824*	−0.1962***	0.0044	−0.0549*
	(0.0196)	(0.0276)	(0.0282)	(0.0329)	(0.0243)	(0.0324)	(0.0330)	(0.0466)	(0.0417)	(0.0245)
Household HIV Unknown	0.0339***	0.0145	0.0496***	0.0738***	−0.0020	0.1742***	−0.1387***	−0.0501*	0.0298**	0.0339
	(0.0085)	(0.0117)	(0.0123)	(0.0157)	(0.0102)	(0.0121)	(0.0146)	(0.0224)	(0.0096)	(0.0188)
Rural	−0.5798***	−0.6052***	−0.5706***			−0.6133***	−0.5093***	−0.5617***	−0.6531***	−0.2153***
	(0.0113)	(0.0159)	(0.0164)			(0.0170)	(0.0191)	(0.0274)	(0.0129)	(0.0246)
Middle Income	0.3122***	0.2939***	0.3297***	0.2199***	0.3771***	0.3626***	0.3699***	0.1314***	0.3332***	0.2884***
	(0.0110)	(0.0151)	(0.0162)	(0.0356)	(0.0121)	(0.0154)	(0.0192)	(0.0345)	(0.0127)	(0.0232)
High Income	0.7344***	0.7063***	0.7725***	0.7855***	0.7262***	0.8510***	0.7427***	0.6059***	0.7561***	0.7654***
	(0.0114)	(0.0157)	(0.0167)	(0.0303)	(0.0131)	(0.0164)	(0.0196)	(0.0318)	(0.0130)	(0.0247)
Prop. of Child. < 5	−0.8006***	−0.1102*	−1.2096***	−1.5137***	−0.5139***	−0.0103	−1.0969***	−2.2341***	−0.6499***	−1.2654***
	(0.0328)	(0.0472)	(0.0469)	(0.0619)	(0.0391)	(0.0475)	(0.0590)	(0.0880)	(0.0381)	(0.0667)
Head Uneducated	−0.8097***	−0.7959***	−0.8360***	−0.8454***	−0.7405***	−0.7688***	−0.8881***	−0.7801***	−0.8366***	−0.6141***
	(0.0100)	(0.0138)	(0.0145)	(0.0183)	(0.0121)	(0.0141)	(0.0171)	(0.0282)	(0.0114)	(0.0212)
Head 40-59	0.2094***	0.1552***	0.2489***	0.2373***	0.1668***	0.0775***	0.2302***	0.5411***	0.2064***	0.1969***
	(0.0102)	(0.0143)	(0.0147)	(0.0184)	(0.0123)	(0.0143)	(0.0191)	(0.0311)	(0.0119)	(0.0204)
Head 60-79	0.3334***	0.2350***	0.4171***	0.3413***	0.2757***	0.1789***	0.2918***	0.5443***	0.2956***	0.3773***
	(0.0127)	(0.0175)	(0.0185)	(0.0237)	(0.0152)	(0.0191)	(0.0221)	(0.0355)	(0.0148)	(0.0257)
Head $\geq$ 80	0.5087***	0.3838***	0.6091***	0.5470***	0.4120***	0.3908***	0.4953***	0.4647***	0.5000***	0.4444***
	(0.0292)	(0.0404)	(0.0424)	(0.0601)	(0.0341)	(0.0442)	(0.0473)	(0.0803)	(0.0346)	(0.0562)
Head Male	−0.2985***	−0.2543***	−0.3554***	−0.2025***	−0.2976***	−0.2053***	−0.3363***	−0.2651***	−0.2984***	−0.2532***
	(0.0101)	(0.0143)	(0.0144)	(0.0177)	(0.0124)	(0.0154)	(0.0170)	(0.0252)	(0.0123)	(0.0185)
Cluster HIV Prevalence	0.0035***	−0.0022**	0.0099***	−0.0076***	0.0093***	0.0266***	−0.0042***	−0.0205***	0.0249***	−0.0014
	(0.0005)	(0.0007)	(0.0008)	(0.0010)	(0.0007)	(0.0009)	(0.0009)	(0.0013)	(0.0025)	(0.0010)
Country Enrollment Rate	0.0439***	0.0412***	0.0469***	0.0237***	0.0513***	0.0520***	0.0437***	0.0124***	0.0425***	0.0434***
	(0.0003)	(0.0004)	(0.0005)	(0.0006)	(0.0004)	(0.0005)	(0.0005)	(0.0008)	(0.0004)	(0.0008)

Table 8: Binary Model